

基因表达调控与生物技术中的酶学

DNA RNA PRO

基因表达调控与 生物技术中的酶学

杨开宇 孟广震 主编
科学出版社

基因表达调控与 生物技术中的酶学

杨开宇 孟广震 主编

406162

科学出版社

1990

内 容 简 介

本书从蛋白质与核酸相互联系的角度论述了酶学的最新进展。全书共有10篇,内容包括基因表达调控的酶学、生物技术中的酶学及分子克隆技术在酶学研究中的应用等。读者既可从本书了解这一领域的研究概貌,也有助于追踪这一领域的新进展。本书可供生物化学、分子生物学及生物工程学的科技人员参考,也可作为大专院校和研究生院有关专业的教材。

27/03/1992

基因表达调控与生物技术中的酶学

杨开宇 孟广展 主编

责任编辑 赵甘泉

科学出版社出版

北京东黄城根北街16号

邮政编码:100707

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1990年7月第一版 开本:787×1092 1/16

1990年7月第一次印刷 印张:13 1/2

印数:0001—1200 字数:313 000

ISBN 7-03-001697-1/Q·250

定价:13.70 元

前言

编写这本书是基于以下的考虑：(1) 和一切生命活动过程一样，基因的表达是通过一系列酶的催化作用实现的，不深入了解相关酶的特性就不可能在本质上阐明基因表达调控的规律及其分子基础；与此相应的，酶是有催化功能的生物大分子，不联系生物功能孤立地研究酶是不全面的，其意义也是有限的。因此，结合生命活动过程研究酶是酶学发展的方向之一。(2) 70年代发展起来的分子克隆技术是一种极其有效的研究工具，其影响所及已超越分子遗传学，对生物学各分支学科（其中也包括酶学）的发展都有极大的促进作用。上面两个特点表明，从蛋白质、核酸两个侧面同时进行研究是当今分子生物学发展的总趋势。为了反映这一趋势在酶学中的表现和已取得的进展，我们特编写本书。其中第1—4, 6, 7篇着重介绍基因表达调控的酶学；第5篇涉及分子克隆中起举足轻重作用的限制性内切酶；而第8—10篇则反映分子克隆技术在酶学中的应用。各篇作者在写作时尽量避免单纯地从酶学或遗传学角度进行介绍，而是力求从二者的结合点上展开。

这本书是由以下8位同志分头执笔编写的：中国科学院上海药物研究所杨胜利和吴汝平同志，中国科学院微生物研究所杨开宇和孟广震同志，中国科学院生物物理研究所王家槐同志，中国科学院上海生物化学研究所施建平同志，中国医学科学院基础医学研究所强伯勤同志及北京师范大学生物系何忠效同志。各位作者的思路、文风及写作格调都不尽相同，而最大的不足莫过于在文章的广度和深度上未必能充分反映这一领域已经取得的全部进展。这是由于编写人员的水平所决定的。改进的办法只有靠不断地补充及同志们的批评指正以至日后的直接参与。从这个意义上讲，这本书只不过是抛砖引玉而已。

在编写过程中我们深深感到，在这一广阔的研究领域中，国人尽管作出过贡献，但毕竟是有限的。这就是差距。弥补的办法是急起直追，迎头赶上。如果这本书在促进我国分子生物学的发展中能发挥点滴作用，这就是我们最大的心愿了。

最后，我们向参加本书编写工作的全体作者表示衷心感谢。由于大家积极、认真而有效率的工作，保证了本书的顺利出版。

主编

1987年5月

目 录

前言

1. DNA 复制及相关蛋白质.....杨胜利(1)
2. 基因转录的分子机制杨开宇(31)
3. RNA 剪接加工及其反应的催化原理.....杨开宇(48)
4. 蛋白质生物合成与氨酰-tRNA 合成酶.....施建平(62)
5. 限制性核酸内切酶强伯勤(83)
6. DNA 甲基化酶及其在基因表达调控中的作用.....何忠效(113)
7. DNA 拓扑异构酶.....王家槐(123)
8. 蛋白质与核酸序列测定技术孟广震(142)
9. 酶的基因工程吴汝平(169)
10. 蛋白质工程及其在酶学研究中的应用.....孟广震、施建平(198)

1. DNA 复制及相关蛋白质

杨 胜 利

1.1 DNA 复 制

1944 年 Avery 首先报道 DNA 可转化肺炎球菌,这是 DNA 可作为遗传物质的第一个重要证据。随后又发现 T₂ 噬菌体蛋白质外壳将 DNA 注射到宿主细胞中可产生许多相同的子代噬菌体,又一次证明 DNA 是遗传物质。1953 年 Watson 和 Crick 根据 X 衍射研究结果提出了 DNA 的互补双螺旋结构模型。1970 年 Crick 又提出中心法则,即遗传信息由 DNA 转录为 RNA,再由 RNA 翻译为蛋白质。这样 DNA 有两个功能:作为遗传信息使细胞具特定表型和自身复制使细胞具有相同的基因型。自 DNA 双螺旋模型提出后的 30 多年来,对 DNA 复制机理及参与复制和复制调控的蛋白质进行了大量研究,阐明了 DNA 复制的一些基本规律和复制蛋白质的功能。Kornberg 编著的《DNA 的复制》(DNA replication)一书详细综述了 1980 年前 DNA 复制机理和复制蛋白质等方面的研究^[1]。

复制是一个半保留的过程,所得的子代分子不仅彼此完全一样,并与亲代 DNA 分子也完全一样。复制通常在 DNA 分子的某一个或几个称为复制起点的特定 DNA 顺序开始,然后单向或双向延伸 DNA 链,双向复制在固定终点终止链延伸。大部分 DNA 分子是双向复制的,单向复制只发生于线粒体和某些质粒。复制时所有的 DNA 链都是由 5' 向 3' 延伸的。大部分 DNA 分子的复制是半连续的,前导链连续延伸,而与之互补的后随链不连续延伸^[1]。

复制可分为三个阶段:复制起始, DNA 链延伸和复制终止。不同阶段需要不同蛋白质。大肠杆菌染色体复制时需要 dnaA 蛋白质、RNA 聚合酶和促旋酶识别复制起点和起始前导链合成。而 DNA 链延伸则需要 DNA 聚合酶 III、单链 DNA 结合蛋白质(SSB)及至少由 7 个蛋白质组成的引物合成复合体(primosome)^[2]。

一种细胞可有多种复制机理。大肠杆菌细胞如缺失了核糖核酸酶 H(RNase H),在缺失复制起点 C(oriC)和 dnaA 蛋白质的情况下,仍可复制染色体,这种方式复制可在没有蛋白质继续合成的情况下正常进行,称为稳定的 DNA 复制,这类复制起始于复制起点 H(oriH)^[3]。

复制的分子生物学和酶学的研究最初集中在 DNA 链的延伸,70 年代后期转向复制起始及其调控的研究,对于复制的终止及子代 DNA 分子的分离研究得较少,近年来对质粒 DNA 分配的研究将有助于阐明终止和分离机理。

1.1.1 DNA 复制起始

复制起始是指参与复制的蛋白质识别复制起点,合成引物,在复制起点形成活性的引

物-模板复合物并按碱基配对加上第一个核苷酸。复制调控主要发生在复制起始阶段。

DNA 聚合酶不能从头合成新生 DNA 链,只能在与模板结合的引物的 3' 末端延伸 DNA 链。大多数复制系统中引物为与模板 DNA 链互补的短 RNA 链,可由 RNA 聚合酶或引物合成酶合成^[4]。但在有些复制系统中引物为蛋白质、tRNA 或四磷酸二腺苷。在 ϕ X174 噬菌体的滚环式复制时,由基因 A 蛋白质作为起始蛋白质在复制,起点产生缺口并与正链形成的共价复合物起始 DNA 合成^[4],而腺病毒由末端蛋白质与模板的共价复合物直接作为引物起始复制的^[5]。反转录病毒由宿主细胞的某一种 tRNA 起始 DNA 合成^[6]。细小病毒基因组为一线性单链 DNA,其末端自身互补,可折回形成一发夹结构,这种次级结构的 3' 末端羟基可起始 DNA 合成^[7]。在爪蛙的卵细胞中 DNA 合成由四磷酸二腺苷起始,三磷酸二腺苷、四磷酸二鸟苷和五磷酸二鸟苷可代替四磷酸二腺苷起始复制^[8]。上述不同起始方式如图 1-1 所示。

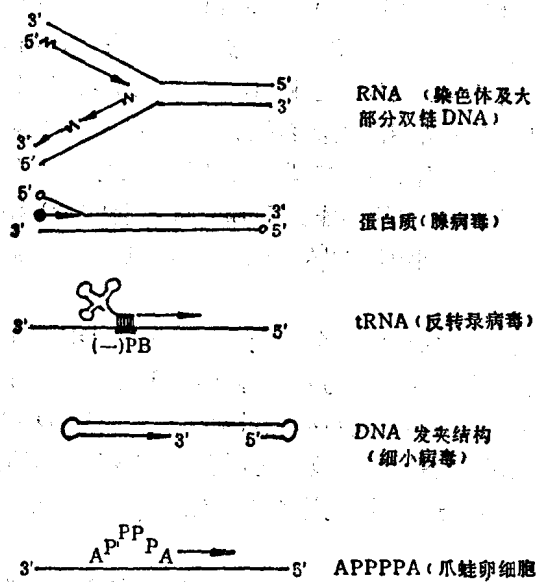


图 1-1 DNA 合成起始引物

W: 引物 RNA;

●: 蛋白质-dGMP 复合物; ✕: tRNA;

(-)PB: 模板上的 tRNA 结合部位。(根据文献[8]改编)

机制。G₄ 噬菌体由引物合成酶合成引物, M13 噬菌体由 RNA 聚合酶合成, 而 ϕ X174 噬菌体则需要一个由 n、n'、n'', dnaB 蛋白质、dnaC 蛋白质和引物合成酶组成的复合物合成引物^[4]。大肠杆菌染色体的引物合成机理与 ϕ X174 噬菌体相似。

双链 DNA 复制是以半连续方式进行的, 前导链的引物在不同细胞中分别由 RNA 聚合酶或引物合成酶合成, 而后随链的引物总是由引物合成酶合成的^[1]。一般引物长度为 10—12 个核苷酸, 而且第一个核苷酸都是嘌呤核苷酸^[12]。

在有些复制系统转录产生引物前体, 需经转录后加工才有起始 DNA 合成的活性。如 ColEI DNA 复制时, 先由 RNA 聚合酶合成一个 RNA 引物前体 RNAII, 再由 RNase H 切除部分 RNA 链产生有活性的引物 RNA, 以起始 DNA 合成^[13]。

RNA 引物起始 DNA 合成是主要的起始方式, 可分为三个阶段, 前起始、引物合成和引物转录后加工。

前起始机制随复制系统而异。大肠杆菌染色体和 λ 噬菌体复制时需转录活化, 即由一在复制起点区外的启动子启动转录并通读至复制起点区, 活化复制起始。转录活化可能是通过形成适合于引物合成的模板结构从而活化复制的^[1]。质粒 pBR322 和 ϕ X174 噬菌体 DNA 复制时, 需先由 n' 蛋白质识别 DNA 链上的效应子位点, 结合于这个位点并活化自身的 ATPase, 然后将参与引物合成的蛋白质在这个位点形成一活性的引物合成复合物, 合成引物^[19-21]。

引物合成可由 RNA 聚合酶或引物合成酶催化。以单链 DNA 噬菌体为模型研究引物合成机理发现有三种不同

1.1.2 DNA 链的延伸

DNA 链的延伸有两种方式。一种是对称的复制,也称为 θ 复制模型。另一种是不对称的复制,即滚环式复制模型(图 1-2)。

Cains 用 ^3H 胸腺嘧啶标记的 DNA 放射自显影图谱研究大肠杆菌染色体的复制中间体结构,并提出了 θ 复制模型。按这个模型新生的 DNA 链随复制叉移动而延伸。但 DNA 聚合酶只能由 5' 向 3' 方向延伸 DNA 链,而双链 DNA 的两条链极性是相反的,一条链由 5' 向 3' 方向延伸,与其互补的链就应由 3' 向 5' 方向延伸。因此复制叉移动时 DNA 聚合酶如何同时复制两条互补的 DNA 链是 DNA 链延伸中的关键。Kornberg 在深入研究 ϕX174 DNA 复制的基础上提出了复制叉移动模型(图 1-3)。按照这个模

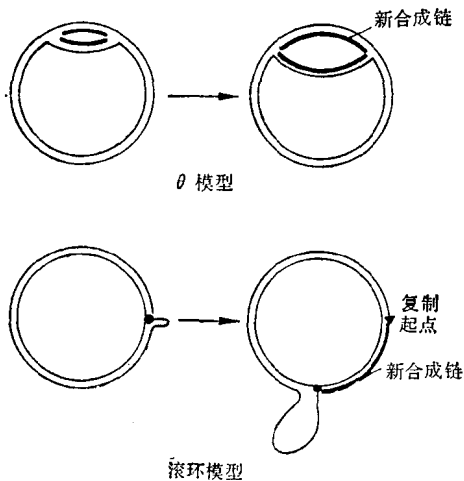


图 1-2 θ 和滚环复制模型(引自文献[1])

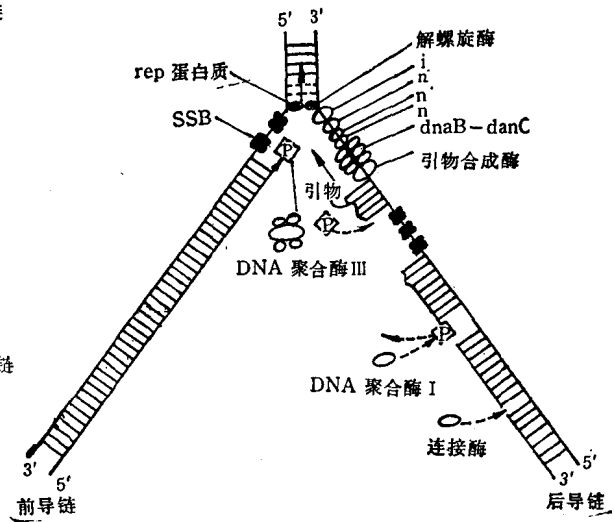


图 1-3 大肠杆菌染色体复制叉(引自文献[35])

型,在复制起点起始复制后,前导链由 5' 向 3' 方向连续延伸。解螺旋酶在复制叉使双链解开,产生局部单链,SSB 与单链 DNA 结合,使单链稳定化,并作为模板合成新生链。这样在复制叉附近的后随链上也出现一段单链区,它可作为不连续复制的模板。由 n 、 n' 、 n'' 和 dnaC 蛋白质等组成的前引物复合体 (preprimosome) 将 dnaB 蛋白质带到这个单链区,并使单链 DNA 构象发生变化,使其适合于引物合成,确立 DNA 合成的起始阶段。然后由引物合成酶合成 RNA 引物,引物在单链区的 5' 末端。再由 DNA 聚合酶 III 在引物 RNA 分子的 3' 末端的羟基起始 DNA 合成。后随链上新生 DNA 链约延伸 1000 个核苷酸左右,延伸方向为 5' 到 3'。最后由 DNA 聚合酶 I 切除引物 RNA,补上 DNA 链,再由连接酶将新生的 DNA 链连接为完整的链。这样随着复制叉的移动,后随链由 3' 向 5' 方向依次出现不连续的单链 DNA 区,而在每一个单链区内是按由 5' 向 3' 方向合成新生 DNA 链,从而完成后随链由 3' 到 5' 的不连续复制。DNA 不对称的滚环复制模型发生于由环状双链 DNA 分子复制单链环状 DNA 分子或线性 DNA 分子。 ϕX174 以这种方式由复制型双链环状 DNA 分子复制单链环状的噬菌体基因组,噬菌体以这种方式复制线性双链 DNA 基因组。 ϕX174 DNA 复制时,先由噬菌体编

码的基因 A 蛋白质在复制型环状双链分子的复制起点切开正链,并和缺口处 DNA 链的 5' 磷酸根共价结合,在 3' 羟基起始 DNA 合成。然后由基因 A 蛋白质和宿主的 rep 蛋白质将双链解开,SSB 与单链 DNA 结合,由 DNA 聚合酶 III 延伸新生 DNA 链。基因 A 蛋白质和 rep 蛋白质不断解开双链,使复制叉沿着负链移动,形成一环状滚环复制,当复制叉到达正链的复制起点时,完成全长基因组合成,由基因 A 蛋白质在复制起点切开正链,开始新一轮滚环复制^[4]。

质粒接合转移时也按滚环模型进行复制,松弛蛋白质的功能可能与基因 A 蛋白质相似^[44]。接合转移的质粒可按 θ 模型和滚环模型两种方式进行复制,这两种复制方式的复制起点是不同的, θ 模型的复制起点称为 ori V,而滚环模型的复制起点称为 ori T。

1.1.3 复制终止和子代 DNA 分配

复制终止包括在终止位点停止 DNA 链延伸,切除 5' 末端的 RNA 引物,填补空隙,然后由连接酶封闭缺口,产生完整 DNA 链。

双链 DNA 复制时,当新生 DNA 链上的引物被切除后在 5' 末端留下一段单链,但没有一个聚合酶可在 5' 末端延伸 DNA 链。因此复制终止中的一个重要问题是完成 5' 末端。DNA 复制时可由下列三种机理完成 5' 末端。

DNA 分子环化 在环状模板上进行复制,最后使新生链的 5' 末端和 3' 末端平置在一起,由连接酶连接为完整环状 DNA 分子。在细菌、病毒和质粒基因组中双链环状 DNA 分子是常见的,而且有些线性双链 DNA 和环状单链 DNA 分子也通过双链环状的复制型进行复制。

末端回文结构 有些线性 DNA 分子末端有回文顺序,可在末端折回形成一自身互补的发夹结构,这样 5' 末端的空隙可由发夹结构的 3' 末端延伸 DNA 链,再由连接酶封闭缺口,然后在回文顺序的 5' 末端由专一的内切酶切开,使发夹结构松开,再由 DNA 聚合酶填补新的空隙,完成 5' 末端。

DNA 分子多联体 T7 噬菌体基因组两端有几百个碱基重复顺序,外切酶 III 切除部分 3' 末端产生 5' 粘性末端,可在分子内或者分子间连接为环状或链状分子,再由 DNA 聚合酶和连接酶完成 5' 末端,使多联体呈共价结合的完整 DNA 链,装配时由内切酶切为单位长度的基因组。 λ 噬菌体也由这种多联体结构完成 5' 末端,然后由终止酶切开多联体,产生单位长度基因组,装配到蛋白质外壳中^[45]。

双向复制时复制终止于一个固定的复制终点, DNA 链延伸至复制终点时,复制叉移动受阻,复制进入终止阶段^[46]。复制终点除了复制终止外还参与子代染色体的分离^[47]。

电镜观察表明细菌染色体 DNA 总是和膜结合的,主要膜结合点是复制起点,复制终点和 DNA 链生长点。新生 DNA 链大部分和膜结合,而复制起点和复制终点可能永久地和膜结合。细胞膜参与子代 DNA 分子向子代细胞分配,它在染色体分配中的作用可能与真核细胞有丝分裂中纺锤体的作用相似。

对子代 DNA 分子的分配机理目前还知道得很少,质粒分配的研究将有助于阐明染色体分配机理。有些质粒的分配位点可直接被宿主的分配机构识别,如质粒 pSC101,但有些质粒分配位点则需与一质粒编码的蛋白质形成复合物后才被宿主分配机构识别,如 F 因子和 P1 质粒。虽然目前还不能详细描述质粒分配的机理,但已有足够证据表明分

配与细胞的分裂,基因重组,SOS 响应性及细胞表层合成有关^[18]。

1.1.4 DNA 复制的保真度

复制的保真度是一种重要的进化力。由于它在衰老、肿瘤形成及疾病中的重要性而愈来愈受重视。DNA 前体、DNA 聚合酶及其他参与复制的蛋白质都与复制的保真度有关。在复制过程中碱基配对,错配的校正及连续程度都会影响保真度。

体外复制表明动物细胞的 DNA 聚合酶由于没有 3'→5' 外切酶的校正功能,复制时保真度较低,误差频率可高达 0.1%。大肠杆菌的 DNA 聚合酶和 T4 噬菌体编码的 DNA 聚合酶具校正功能,复制误差频率为 0.001%。大肠杆菌 DNA 聚合酶 I 复制保真度比 DNA 聚合酶 III 低,主要是由于合成 DNA 的转换率较低,仅为 0.7—1.7%,而 DNA 聚合酶 III 的转换率高达 6—13%^[19]。

细胞内有各种修复系统参与 DNA 损伤的修复及错配碱基的校正。大肠杆菌体内至少有三种修复机理:直接修复、切口一切除一聚合修复和倾向错配修复。直接修复发生在复制过程,而后两种修复机理发生于复制后,这里将只讨论直接的修复机理。

大肠杆菌中有两种直接修复机理:校正和甲基指导的修复。原核 DNA 聚合酶都有 3'→5' 外切酶活力,可识别和切除不配对的 3' 末端,可在聚合反应中,每加上一个核苷酸前检查一下前一个核苷酸配对是否正确。在引物的 3' 末端如有一段不配对的核苷酸链,3'→5' 外切酶可依次切除不配对的核苷酸,然后在配对核苷酸的 3' 末端起始 DNA 合成,提高了复制的保真度。突变的 T4 DNA 聚合酶的校正功能与酶的外切酶活力与聚合酶活力的比值有关,外切酶比活性高,校正功能也较强,自发突变频率较低^[20]。

当错误参入的不配对核苷酸在聚合反应中未被校正可通过甲基指导的修复系统进行修复。这个修复系统由 DNA 链上核苷酸的甲基化控制。大肠杆菌的 DNA 中有少量甲基胞嘧啶和甲基腺嘌呤,他们可保护 DNA 链不受限制酶的攻击。复制时新生的 DNA 链暂时未被甲基化酶修饰,而模板链已甲基化。当发生核苷酸错配时,限制酶可由甲基指导选择性地识别新合成链上的错配核苷酸,以提高复制保真度。在甲基指导的修复机理中,细胞内甲基化酶的活力是十分临界的,甲基化酶活力要维持在正好使新合成的 DNA 链暂时未甲基化。甲基化酶活力过高或过低可导致新合成链过快甲基化或模板链甲基化不完全,从而降低复制保真度^[20]。

1.2 参与复制和复制调控的蛋白质

DNA 复制系统比转录系统复杂,一般都由多种蛋白质组成,DNA 聚合酶只是其中一类。 ϕ X174 噬菌体复制至少需要 13 种宿主蛋白质和 1 种噬菌体编码的蛋白质^[21,22]。T4 噬菌体复制除了需要宿主的复制蛋白质外还需要 7 种噬菌体编码的蛋白质^[23]。已报道的大肠杆菌中参与染色体复制和复制调控的蛋白质至少有 19 种(表 1-1)。他们可分为三类,起始蛋白质、复制蛋白质和专一性蛋白质。起始蛋白质参与转录活化,复制起点识别,前导链引物合成。复制蛋白质参与双链解开,后随链引物合成,DNA 链延伸,引物切除,空隙填补,新合成链连接。专一性蛋白质抑制非专一性的复制起始,使 DNA 复制起始于特定的复制起点。

表 1-1 参与大肠杆菌染色体复制的蛋白质*

蛋 白 质	功 能
DNA 聚合酶 I	切除引物 RNA, 完成 5' 末端, 错配校正
DNA 聚合酶 III	DNA 链延伸, 错配校正
RNA 聚合酶	转录活化, 引物 RNA 合成
dnaA 蛋白质	转录活化, 复制起始
dnaB 蛋白质	引物 RNA 合成, ATPase
dnaC 蛋白质	引物 RNA 合成
引物合成酶	引物 RNA 合成
SSB	复制起始, DNA 链延伸
rep 蛋白质	解链, 复制叉移动
解螺旋酶 II	解链, 复制叉移动
拓扑异构酶 I	专一性因子, 拓扑异构
促旋酶	拓扑异构产生负超螺旋, 复制起始
HU 蛋白质	复制起始, 专一性因子
RNase H	专一性因子抑制 oriH 复制起始, 切除引物 RNA
连接酶	连接 DNA 片段
i 蛋白质	引物 RNA 合成
n 蛋白质	引物 RNA 合成
n' 蛋白质	引物 RNA 合成, ATPase
n'' 蛋白质	引物 RNA 合成

* 根据文献 [1]、[35] 改编。

1.2.1 DNA 聚合酶与反转录酶^[1,24]

DNA 聚合酶以 DNA 链为模板, 按碱基配对逐个聚合脱氧核苷酸, 合成与模板互补的 DNA 链。它催化引物或 DNA 链的 3' 末端羟基与新加核苷酸的 α -磷酸根生成磷酸二酯键。催化的反应是可逆的。DNA 聚合酶不能从头 (*de novo*) 合成 DNA, 只能从引物的 3' 末端延伸 DNA 链, 链延伸的极性总是从 5' 到 3'。

不同生物的 DNA 聚合酶和同一种细胞中不同的 DNA 聚合酶结构和功能各不相同, 主要表现在下列几方面。

酶的组成和大小 大肠杆菌的 DNA 聚合酶 I 是一条分子量为 109kd 的单一肽链, 而 DNA 聚合酶 III 全酶是由 7 种亚基组成分子量高达 382kd 的复合物。

模板与引物的选择性 大肠杆菌的 DNA 聚合酶 I 可在带引物的单链或带缺口的双链合成 DNA, 而 DNA 聚合酶 II 和 III 在带有小于 100 个核苷酸单链间隙的双链 DNA 模板上合成 DNA。

保真度 原核 DNA 聚合酶具 3'→5' 外切酶活力, 复制保真度较高, 而真核 DNA 聚合酶无 3'→5' 外切酶活力, 复制保真度较低。原核细胞的 DNA 聚合酶由于外切酶活力与聚合酶活力的比值不同, 复制保真度也各不相同。

复制连续程度 DNA 聚合酶合成 DNA 时有两种方式, 连续型和不连续型。当 DNA 聚合酶共价连接一个核苷酸后仍与模板结合, 直至 DNA 链合成终止, 这称为连续型的。相反 DNA 聚合酶每共价连接一核苷酸后就与模板解离, 与游离的酶平衡, 然后随机地参与下一个核苷酸的聚合, 则称为不连续型的。DNA 聚合酶合成 DNA 链时连

续程度取决于 DNA 聚合酶的本性与状态,模板和引物的性质,前体核苷酸的浓度及聚合反应的条件等因素。

催化效率 体外测定各种 DNA 聚合酶活力在每摩尔酶每分钟聚合几百到几千个核苷酸之间。

缺口转移和链取代合成的能力 催化这两种反应依赖于 DNA 聚合酶的 5'→3' 外切酶活力。DNA 聚合酶 I 具有缺口转移和链取代合成能力。

生理功能 不同的 DNA 聚合酶可分别参与复制、修复或重组,有的 DNA 聚合酶可具多种生理功能。

1.2.1.1 原核细胞的 DNA 聚合酶

大肠杆菌有三种 DNA 聚合酶,按其发现次序分别称为 DNA 聚合酶 I、II 和 III。这三种酶都具有 3'→5' 外切酶活力,即具有校正功能。

DNA 聚合酶 I 为一分子量为 109 000 的单一肽链,呈球状,直径约为 65 Å。整个分子折叠为两个区域,由一个对蛋白酶敏感的肽键相连接。用胰蛋白酶水解可得两个片段,大片段称为 Klenow 片段,具聚合酶和 3'→5' 外切酶活力,而小片段具 5'→3' 外切酶活力。DNA 聚合酶可催化三种聚合反应:空隙填补 (gap-filling)、缺口转移和链取代合成。后两种反应是 DNA 聚合酶 I 所特有的,依赖于它的 5'→3' 外切酶活力,可在缺口的 5' 末端开始降解双链 DNA,也可切除 DNA-RNA 杂合分子中的 RNA 链,在复制中可切除引物 RNA 链。

DNA 聚合酶 I 参与复制和修复,它的缺口转移功能在修复中起关键作用。在复制中 DNA 聚合酶 I 参与不连续复制,它的功能是切除引物 RNA 链和补上 DNA 链。

DNA 聚合酶 I 催化聚合反应时先与 DNA 模板和 dNTP 形成一复合物,然后发生构象变化,转化为活化型复合物,构象变化是 DNA 聚合酶 I 催化聚合反应的速度限制因子^[25]。

DNA 聚合酶 II 的结构和功能还不太清楚,它的缺失突变株对 UV 的敏感程度及转导和接合转移时的重组频率都无明显变化,而且也未发现细胞的其他缺陷。

DNA 聚合酶 III 是大肠杆菌的主要复制酶,它的结构与 RNA 聚合酶相似,由多种亚基组成,各种亚基的分子量,编码基因和功能见表 1-2。

表 1-2 DNA 聚合酶 III 全酶: 亚基和功能*

亚 基	分子量 (kd)	基 因	功 能
pol III* pol III' pol III α ε θ τ factor II γ δ Co pol III* β	140	dnaE	聚合酶
	27	dnaQ	3'→5' 外切酶
	10	?	?
	78	dnaZ	形成二聚体
	52	dnaX	将 β 带到 DNA 模板
	32	dnaX-Z	将 β 带到 DNA 模板
	37	dnaN	聚合起始及连续度

* 根据文献 [26] 改编。

α、ε 和 θ 三个亚基构成核心酶,称为 pol III, pol III 在双链 DNA 的单链空隙有有

限催化活力, ϵ 亚基具 $3' \rightarrow 5'$ 外切酶活力^[26]。pol III 和 τ 亚基的复合物称为 pol III', 它与核心酶最大结构差异是以二聚体存在。 τ 亚基有依赖于单链 DNA 的 ATPase 活性, 它的功能可能是使两个 pol III 分子结合在一起。在 pol III' 上再加 δ 和 γ 亚基得 pol III*, δ 和 γ 亚基的功能是将 β 亚基运送到 DNA 模板。 β 亚基在全酶中以二聚体存在, 它对于复制起始和连续度是必需的。聚合酶 III 具 $3' \rightarrow 5'$ 外切酶和 $5' \rightarrow 3'$ 外切酶活性。但它的 $5' \rightarrow 3'$ 外切酶活力与 DNA 聚合酶 I 不同, 对双链 DNA 无活力, 只作用于单链 DNA, 因此 DNA 聚合酶 III 虽有 $5' \rightarrow 3'$ 外切酶活性, 但无缺口转移和链取代功能。

DNA 聚合酶 III 有两种, 一种可在 ATP_{rs} 存在时形成起始复合物, 而另一种只能与 ATP 形成起始复合物, 当 ATP_{rs} 存在时起始复合物水解。进一步研究发现这不是二种 DNA 聚合酶 III 分子, 而是同一全酶二聚体的两个单体。可能是其中一个单体多一个亚基或者一个亚基受共价修饰, 也可能是由别位相互作用引起的结构差异。这种结构的不对称性对于酶的功能是十分重要的, 使它在复制又能同时催化两条链的不对称合成, 即同时催化前导链和后随链的合成, 而这两条链的复制机制是不同的^[27]。在多引物的模板上, DNA 聚合酶 III 与模板结合后, 沿着 DNA 链迅速移动, 与最先遇到的引物形成起始复合物, 起始 DNA 合成^[28]。

大肠杆菌的 T 系噬菌体编码自己的 DNA 聚合酶, 这些酶都有 $3' \rightarrow 5'$ 外切酶活力, 而无 $5' \rightarrow 3'$ 外切酶活力。T7 噬菌体编码的 DNA 聚合酶结构与其他 DNA 聚合酶不同, 由两个亚基组成, 一个是噬菌体编码的 T7 的基因 5 蛋白质和宿主的 trx A' 的硫氧还蛋白 (thioredoxin)。

其他原核细胞也具有与大肠杆菌 DNA 聚合酶 I、II 和 III 相应的结构和功能相似的三种 DNA 聚合酶。

1.2.1.2 真核细胞的 DNA 聚合酶^[1,29]

动物细胞中已发现三种 DNA 聚合酶, α 、 β 和 γ 。DNA 聚合酶 α 在真核细胞中普遍存在, 是真核细胞的主要复制酶, 存在于细胞核和细胞质中, 在生长细胞中占 DNA 聚合酶活力的 85% 以上, 而在休止细胞中仅占 5%, 表明在染色体复制期大量合成 DNA 聚合酶 α 。与大肠杆菌 DNA 聚合酶 III 相似, DNA 聚合酶 α 也是由多种亚基组成, 核心酶与全酶的结构与功能都不相同。一些附加蛋白质对 DNA 聚合酶 α 的活性有很大影响, 可调节它与模板-引物复合物的相互作用, 聚合反应的速度以及复制保真度。解螺旋酶和依赖于 DNA 的 ATPase 均可提高 DNA 聚合酶 α 的活力。在纯化 DNA 聚合酶 α 全酶时, 得到 4 种类型的复合物, A 型复合物中除 DNA 聚合酶 α 外还有依赖于 DNA 的 ATPase, $3' \rightarrow 5'$ 外切酶和拓扑异构酶 II, A 型复合物可能参与前导链的复制。B、C 和 D 型复合物都有引物合成酶可能参与后随链合成。DNA 聚合酶 α 在复制叉催化聚合反应的机理可能与 DNA 聚合酶 III 相似, 以一不对称的二聚体分子同时催化前导链和后随链的复制^[30]。

DNA 聚合酶 β 存在于细胞核中, 在细胞中的浓度相对恒定。淋巴细胞中在 DNA 修复活性最高时, DNA 聚合酶 β 的活力达最高值。照射可使大白鼠神经元细胞的 DNA 聚合酶 β 活力升高 10 倍左右。它可能参与修复和重组。

DNA 聚合酶 γ 存在于线粒体和核中,参与线粒体,腺病毒及细小病毒 DNA 复制,它参与的 DNA 复制中都有一个链取代合成阶段,因此可能在链取代合成中起关键作用。在三种真核 DNA 聚合酶中只有 γ 能以连续方式合成 DNA 链。

单细胞真核细胞的 DNA 聚合酶与动物细胞的 DNA 聚合酶不同,而与原核 DNA 聚合酶相似,具有 $3' \rightarrow 5'$ 外切酶活性,因此有校正功能。牛痘病毒编码的 DNA 聚合酶也具有 $3' \rightarrow 5'$ 外切酶活性。

1.2.1.3 反转录酶

反转录酶最初在白血病病毒和 Rous 肉瘤病毒中发现,在反转录病毒复制中起关键作用,参与由病毒 RNA 反转录 DNA 和 DNA 复制。它具有 DNA 聚合酶和 RNase H 活力,可以 DNA 或者 RNA 为模板合成 DNA,合成极性与其他 DNA 聚合酶一样也是由 $5' \rightarrow 3'$ 。

反转录酶由病毒 RNA 拷贝 DNA 时以 RNA 为模板, tRNA 为引物。引物 tRNA 与病毒 RNA 链的(一) PB 位点共价结合,然后由反转录酶在 tRNA 3' 末端起始 DNA 合成。所得的 DNA-RNA 杂合分子由反转录酶的 RNase H 活力水解 RNA 链得单链 DNA,再由反转录酶合成互补的 DNA 链。

一些植物的 DNA 病毒如花菜花叶病毒的染色体复制时需要反转录酶。病毒 DNA 作为复制和转录的模板先由 RNA 聚合酶 II 转录产生 35S RNA,然后 35S RNA 与 met tRNA 结合起始 DNA 合成,聚合反应由反转录酶催化,所得杂合分子中的 RNA 链由反转录酶中的 RNase H 切除,最后由 DNA 聚合酶合成互补链,得双链 DNA 病毒基因组。

成骨髓细胞瘤病毒的反转录酶需经翻译后加工才有活性。由二个 β 亚基组成的反转录酶前体经蛋白水解酶加工得全酶 $\alpha\beta$, α 亚基是 β 亚基由蛋白水解酶切除 30kd 所得 65 kd 肽链,具聚合酶和 RNase H 活性,在病毒颗粒中也存在游离的 α 亚基,但 $\alpha\beta$ 复合物对于与模板有效地结合是必需的。反转录酶活力受磷酸化调节,在体外用蛋白质激酶处理可提高反转录酶的聚合酶活力,用磷酸酯酶处理则降低其聚合酶活力。

1.2.1.4 RNA 复制酶

RNA 复制酶存在于正链 RNA 噬菌体中,这类噬菌体的基因组是单链 RNA,由 RNA 复制酶直接拷贝 RNA,不经 DNA 进行复制。RNA 复制酶是依赖于 RNA 的 RNA 聚合酶,既不同于 DNA 聚合酶也不同于 RNA 聚合酶。它由 $5'$ 向 $3'$ 方向延伸 RNA 链,复制起点总是 $5'$ 末端的 GTP。

$Q\beta$ 噬菌体的 RNA 复制酶由四个亚基组成,其中亚基 II 由噬菌体编码,其他三个亚基为宿主蛋白质,它们是核糖体蛋白质 S1 和蛋白质合成延伸因子 Tu 和 Ts。f2 噬菌体复制酶的三个宿主编码的亚基与 $Q\beta$ 噬菌体复制酶一样。延伸因子 Tu 和 Ts 对于 $Q\beta$ 复制酶中基因 II 亚基维持活化型构象是必需的。 $Q\beta$ 复制酶没有校正功能,复制时产生突变的频率为 10^{-3} 到 10^{-4} 。 $Q\beta$ 复制酶还可不依赖模板随机合成 RNA^[31]。

1.2.2 引物合成酶与 RNA 聚合酶

RNA 聚合酶与 DNA 聚合酶的一个重要区别是 DNA 聚合酶不能从头开始合成,

只能在引物的 3' 末端起始 DNA 链合成。在大多数细胞中引物为 RNA。它与模板互补,可由引物合成酶或 RNA 聚合酶合成,后随链不连续复制的引物都是由引物合成酶合成的。

引物合成酶广泛分布于各种生物体,在复制中起重要作用。大肠杆菌的引物合成酶由染色体基因 *dnaG* 编码,分子量为 60kd,参与染色体和某些质粒后随链引物合成和一些单链噬菌体复制的引物合成,它对于质粒 pSC101 稳定的维持是必需的^[43]。大肠杆菌引物合成酶可催化核糖核苷酸和脱氧核糖核苷酸聚合,所合成引物一般为 10—60 个核苷酸。大肠杆菌引物合成酶合成引物还需其他蛋白质,所需蛋白质随模板而异,现已发现三种引物合成系统。最简单的引物合成系统仅由引物合成酶和 SSB 组成,可有效起始 *G*₁、ST-1、 α -3 及 K 等单链 DNA 噬菌体单链模板引物的合成,所合成引物在 26 到 29 个核苷酸之间^[4]。第二个引物合成系统由引物合成酶与 *dnaB* 蛋白质组成,当单链 DNA 没有和 SSB 结合时可由这个系统合成引物,所试的天然单链 DNA 都可作模板,所合成的引物在 10 到 60 个核苷酸。这个系统可在几个位点起始引物合成,由于它不是在专一位点起始引物合成,称为一般引物合成系统。另一个引物合成系统较复杂,由引物合成酶、*dnaB*、*dnaC*、*n*、*n'*、*n''*、*i* 等 7 种蛋白质组成,它参与 ϕ X174 和大肠杆菌染色体的引物合成。由于它在专一位点起始引物合成,因此称为专一的引物合成系统^[33]。*i* 蛋白质除了参与引物合成之外,还参与稳定 DNA 复制的诱导^[44]。*n'* 蛋白质是一个依赖于单链 DNA 的 ATPase,识别复制起点,决定了模板及复制起点的专一性,它与复制起点附近的效应子位点结合后,使引物合成系统在效应子位点正确装配,并水解 ATP 提供能量,以供引物合成蛋白质复合物沿模板移动用,而 *dnaB* 蛋白质使模板 DNA 发生构象变化以适合引物合成^[35]。这个系统合成引物一般为 10—12 个核苷酸。

除了引物合成酶外 RNA 聚合酶也可合成引物, RNA 聚合酶可在 M13, f1, fd 等单链 DNA 模板上的特定位点起始引物合成,合成引物约 30 个核苷酸。RNA 聚合酶还参与质粒 Col EI 前导链引物合成。

一些大肠杆菌噬菌体和质粒编码自己的引物合成酶,如 T4、T7 噬菌体及 B、C、I₁、I₂、K、M、P 和 U 不相容族的质粒都编码自己的引物合成酶,他们与大肠杆菌 *dnaG* 引物合成酶不同,只能以核糖核苷酸为底物,不能合成脱氧核苷酸链。第一个发现有编码引物酶能力的质粒是 R64drd-11,由于编码引物酶的基因可抑制 *dnaG* 的突变,故称为 *sog* 基因^[46]。随后在 RP4 和 R16 质粒中也发现引物合成酶基因,他们除了参与后随链引物合成外,还参与质粒接合转移^[37,38]。在接合转移中引物合成酶有两个作用,作用于供体细胞促进 DNA 合成以取代转移的链和在受体细胞中合成引物,用以合成互补链^[39]。T7 噬菌体的基因 4 蛋白质和 T4 噬菌体的基因 41/61 蛋白质也具引物合成酶能力,T 系噬菌体的引物合成酶在噬菌体模板上合成的引物较短,仅 4—5 个核苷酸^[31]。

RNA 聚合酶在 DNA 复制中有两种功能,合成前导链的引物,如在 Col EI 质粒系统,和转录活化,如在大肠杆菌染色体和 λ 噬菌体系统。

1982 年相继在 Ehrlich 腹水瘤和果蝇胚胎细胞中发现引物合成酶^[40,41]。最近发现人的线粒体引物酶与 5.8S 的 rRNA 结合在一起,这个 rRNA 对于引物合成酶的活力是必需的^[42]。

1.2.3 核糖核酸酶 H (RNase H)^[43]

RNase H 专一地水解 DNA-RNA 杂合分子中的 RNA 链,产生 2—9 个核苷酸的寡核苷酸链。RNase H 在生物体内普遍存在,1969 年首先在大肠杆菌中发现 RNase H,1978 年 Tomizawa 等报道 RNase H 对于 Col E1 DNA 复制是必需的^[44]。但随后又发现在体内 RNase H 对于 Col E1 DNA 复制不是必需的,在 RNase H 缺失的细胞中质粒 Col E1 仍能复制^[45]。Col E1 DNA 能以同样频率转化 rnh^+ 和 rnh^- 宿主。Col E1 DNA 复制依赖于 DNA 聚合酶 I,在 pol A 宿主中不能增殖, rnh 突变可抑制 pol A 突变,在 rnh 和 pol A 双突变的宿主中 Col E1 DNA 可正常复制,质粒拷贝数与在野生型宿主相同。一个质粒 Col E1 的复制起点变种 *cer6* 在 rnh^+ 和 pol A⁺ 宿主中都不能复制,但在 rnh^- 宿主中可复制。RNase H 在 Col E1 DNA 复制中参与引物的加工,使引物 RNA 前体转化为有起始活性的引物。上述结果表明在大肠杆菌野生型宿主中先由 RNA 聚合酶转录产生引物前体,然后由 RNase H 在特定位点切断前体 RNA,产生活性引物,再由 DNA 聚合酶 I 在引物的 3' 末端起始并延伸 DNA 链。这是 Col E1 DNA 的首选复制机理。但在这个首选复制起点外,Col E1 DNA 还有其他复制起点,RNase H 抑制了这些复制起点的复制,使复制在专一复制起点,以依赖于 DNA 聚合酶 I 的复制机理进行。在 rnh^- 宿主中其他复制起点的复制起始抑制解除,而这些复制起点起始的复制不依赖于 DNA 聚合酶 I,所以在 rnh^- 和 pol A 双突变宿主中质粒 Col E1 能正常复制^[43]。

RNase H 在大肠杆菌染色体复制中的作用与在 Col E1 DNA 复制中相似。在野生型大肠杆菌中染色体复制专一地起始于 *ori C*,复制起始依赖于 *dna A* 蛋白质。*dna A* 或 *ori C* 缺失对于细胞是致死的,氯霉素是一种蛋白质合成抑制剂,它与 *dna A* 突变一样,可阻断大肠杆菌染色体复制起始。Kogoma 等分得一株 *sdr A⁻* 突变株,它的染色体复制对氯霉素呈抗性,在这个变种中 *ori C* 缺失或 *dna A* 突变对染色体复制没有影响。这表明在 *sdr A⁻* 宿主中染色体复制机理与野生型细胞不同。基因定位与表型研究表明 *sdr A* 与 *rnh* 为同一基因,编码 RNase H。在 rnh^- 宿主中染色体复制起始于一系列称为 *ori H* 的复制起点。上述结果表明 RNase H 抑制了 *ori H* 的复制起始,使染色体复制专一地起始于 *ori C*^[43,46]。

除了在质粒 Col E1 和大肠杆菌染色体复制中 RNase H 起专一性因子作用外,在有些复制系统中 RNase H 参与引物切除或 RNA 模板链切除,如在反转录病毒复制系统。

1.2.4 DNA 连接酶

在 DNA 连接酶发现以前,DNA 复制、DNA 修复和基因重组研究都发现需要一个酶连接开裂的 DNA 链骨架,1967 年五个实验室先后报道发现了 DNA 连接酶。

DNA 连接酶在生物体内广泛存在,它催化相邻核苷酸的 3' 羟基和 5' 磷酸根之间形成磷酸二酯键,DNA 连接酶催化的反应需辅酶 ATP 或 NAD,反应分两步进行,先将辅酶的腺苷酰基转到 DNA 连接酶的一个赖氨酸残基的 ϵ 氨基上,然后腺苷酰基由酶转到 DNA 链缺口的 5' 磷酸根上,最后在缺口的 3' 羟基和活化的 5' 磷酸根之间形成磷酸

二酯键,放出 AMP。

DNA 连接酶与 RNA 连接酶不同, DNA 连接酶只能连接双链,而 RNA 连接酶可连接单链。不同的 DNA 连接酶性质也不同,如大肠杆菌与 T4 噬菌体连接酶的辅酶与底物专一性均有所不同。大肠杆菌连接酶需辅酶 NAD,只能连接带粘性末端的双链 DNA 分子,而 T4 噬菌体连接酶则需 ATP 作为辅酶,可连接带粘性末端或平端的双链 DNA,而且可连接 DNA-RNA 杂合分子。RNA 连接酶可使 T4 DNA 连接酶连接平端双链 DNA 的活力增高 20 倍,它并不增高酶与底物的亲和力。

大肠杆菌 DNA 连接酶由位于染色体图 52 份的 lig 基因编码,lig 基因已克隆和测定了核苷酸顺序,并由此推出氨基酸顺序^[47]。大肠杆菌连接酶由 671 个氨基酸组成,沉降系数较低,为 3.9S。连接酶对蛋白酶比较敏感,水解产物催化形成酶-AMP 复合物的活力不变,但将 AMP 转移到 DNA 并形成磷酸二酯键的能力丧失。DNA 连接酶对于 DNA 复制是必需的,连接酶缺失的细胞不能存活。在 DNA 复制中连接酶参与后随链复制的冈田片段连接,形成完整的后随链的互补链。

真核细胞的 DNA 连接酶催化机制与原核细胞 DNA 连接酶相同,也需腺苷酰基激活,参与反应的辅酶为 ATP。哺乳动物细胞中至少有两种抗原无关的连接酶,分别称为 DNA 连接酶 I 和 II。连接酶 I 分子量为 200 kd,而连接酶 II 为 85 kd。在增殖细胞中 DNA 连接酶 I 活力占主要地位,而在休止细胞中主要是连接酶 II。

1.2.5 单链 DNA 结合蛋白质 (SSB)

在细胞内 DNA 重组只需几分钟,重组中互补 DNA 链的复性是速度限制步骤。在体外即使在复性最适溶液中 DNA 也不发生复性,这是由于单链 DNA 中的局部发夹式次级结构妨碍了复性,可提高复性温度或用变性剂使发夹结构破坏。在细胞内是由 SSB 熔化发夹结构,使复性能在 37℃ 进行,促进链重组。SSB 不仅在重组中起重要作用,而且对于复制也是必需的。

第一个发现的 SSB 是 T4 噬菌体的基因 32 蛋白质。后来在各种原核和真核细胞中都发现了 SSB。SSB 可在远远低于 DNA T_m 的温度转化 ds DNA 为 ss DNA,SSB 通过与单链 DNA 紧密结合使 DNA 熔化,它被先后称为解链蛋白质,螺旋去稳定蛋白质和单链 DNA 结合蛋白质。SSB 转化双链 DNA 为单链 DNA,破坏单链 DNA 中二级结构,这又有助于 DNA 复性,因此 T4 的基因 32 蛋白质又被称为复性蛋白质。功能的多样性造成了这类蛋白质命名的复杂化,但不管它的功能是什么,与单链 DNA 紧密结合是他们共同特点,因此本书中把这类蛋白质称为单链 DNA 结合蛋白质。

SSB 优先与双链 DNA 中不稳定的区域结合,用 SSB 得到的 λ 噬菌体的变性图与用碱或高温处理所得的变性图相似。SSB 与单链 DNA 结合的一个特征是高度协同,SSB 分子一个接一个排列在 DNA 单链上,当单链 DNA 与三分之一饱和量的 SSB 混和时,结果 SSB 集中在三分之一区域,而其他区域几乎没有 SSB。

T4 噬菌体基因 32 蛋白质对于噬菌体基因组的复制和重组都是必需的。在整个复制过程中都需要基因 32 蛋白质。SSB 与酶不同,不具催化活性,它同时改变了反应的速率和平衡状态。在复制过程中需新生 DNA 链化学计量量的基因 32 蛋白质,每个复制叉约需 170 个分子。在 T4 噬菌体感染的大肠杆菌细胞中约有 60 个复制叉,因此约需一